RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/549,943
Source: IFW0
Date Processed by STIC: 04/05/2007

ENTERED

IFWO

RAW SEQUENCE LISTING DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:22

Input Set : **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

<110> **APPLICANT:** Genencor International, Inc.
Jones, Brian E.
Grant, William D.
Heaphy, Shaun
Grant, Susan
<120> **TITLE OF INVENTION:** Novel Bacillus 029cel Cellulase
<130> **FILE REFERENCE:** GC796-2-PCT
C--> <140> **CURRENT APPLICATION NUMBER: US/10/549,943**
C--> <141> **CURRENT FILING DATE: 2005-09-20**
<150> PRIOR APPLICATION NUMBER: US 60/466,831
<151> PRIOR FILING DATE: 2003-04-29
<160> NUMBER OF SEQ ID NOS: 3
<170> SOFTWARE: FastSEQ for Windows Version 4.0
<210> SEQ ID NO: 1
<211> LENGTH: 3410
<212> TYPE: DNA
<213> ORGANISM: Bacillus sp.
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(3410)
<223> OTHER INFORMATION: isolated from environmental sample from Sonachi Lake, Kenya
<400> SEQUENCE: 1

atcaacacgc tggaaagtaa tttcaagggt aaggccatcg gttgccgccg gggtagaaat 60
gtgcggttgg atttcgttga gcggcgtcgc cggcgttcca ccgagggcat agcgcagcag 120
gttggcgatg ccaccggtga ggccttcggg gccgcctacg atgttgtgct cagccgccca 180
tgcgatgtag ccgtccggct cgggttcgct cgcgggggtg aagaagacaa tgtcgtcgag 240
ataaaggttg ccgcttccgc tctcaacgcc gccgaggttg aattggattt cgcaaattct 300
cgttaggtcc agcacggaat cgccgacgag gtcggctatg ggaatctgaa tgcgcccata 360
gggttgggta cgcggaaggg acacgtaggg acccactttg tcattgggcg agacgagccg 420
gacaaagatt tggtgcgccg cctgcgaggg gccttggagg gcgagagaaa ggtacgtgag 480
ggcgctgatg tcgtgcgtgg gaccgtctcc ccagttgtcg agattgagcc caaatccggc 540
ccaccatccg gcgatagtgt agctccaatg gtagtgacgc tcaccctcga agccgccgct 600
ggagagttcc tgcaagccgt cgccccaaat gcccgtgatg agcgttgcct cgtcacggta 660
gatcacaagt tcggcggcgg gtgccggggg aagatcgcct tgagtgatca cgagagtggc 720
ggtggcgctg ccttcgtgat tagggtcggt aatggtggcg acgaccgtgt agctaccggg 780
ccccactggc gcatgggtgg aaccgttgta ggtaaaggag acgtcaagcc ccacgggatg 840
ggtctcggca agagcggcct tgggggtgcc gtcgaaaacg tgttccaaat tggagagcgt 900
gatggtggcg ggtgccttga gcacagtcac agaaacagtg gattgcacgg gatcgtgcgc 960
tgccgtgtct gcaggtgtga agaccacgct gtaaaaacgg gttccggcgg acggtgcaag 1020
gccggacagg acaaaggcaa agtcgccggg gacggcggct actccgccgc tcaggccggc 1080
ctccgcaagg gtttgcccga aggtgatggg tgcggctgtg ggccacatct ccacaaggcc 1140
ggtgtccccc tcgtcacgca ccggcatgag ggcggagagg agatgaatgt aactggcttg 1200
gtaattgatg tcgggctcgg tgatttccca tgagttctcc ggccaaaaac cattccaatc 1260

RAW SEQUENCE LISTING DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:22

Input Set: **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

aaggtaggct ttttgcacgg gttggtctcg gatcgcctga atgcttccgc tgtatttggg 1320
cattgggacc cgcccgaaag aaaaccagga gcgggaccgt agagtgaagt gagggcattg 1380
tcccagtccg gccatcgcgg aaccaatggt ggtagatttc attggctgca cggtcagcgc 1440
cgctggcata catgttgcta agatagacca tgcccattgg gttcactccg tggagatagt 1500
gcaggtagcc catcgcggca tcgcgatgcg cggccgcgtc ggcggggttg agcccaagcc 1560
tccgtacccc ctcgaagaaa aagccagcct gagactttgt tttgttcgag ccccacgtgt 1620
aatcctgatc cttcaggtag gcgcggtagg cgtcggtctg gttattccat gcaccgagaa 1680
actccccacc gtttatagaa gccgccatcc ggttgcggat gtcggcagag acgctaggcg 1740
tcgctcccgg gagggtcgtg tagtgggcga gagcttttg tagctcacct tgaaagggga 1800
agaaatacca ccactgcacg ggctccatat cgagatagcg cacatcgaag aaatcgcgat 1860
agaccgcacc gcccgtgcgc tcgaagagca tggcggcggc catcacacgg ttggctagcg 1920
tatcgtgggc attgcgcgag gggctcacgg aagcaaatcc ggtgttgtcg aaaggcacat 1980
gaggatggac catggtccaa ttccatgcgg cgatggcagc ggattcgagg gtgacggcat 2040
aatcgctcat gcctacgctc tcaaagacag tcgccccgag ggcgaaagcg gcggcagcca 2100
tggcagtggc ctcggtcgag acggggccgt agtaacgcgg atgggtgtcg gtgctcggcg 2160
ggctggcgct ctggtgcccc gtcacggaaa ctttcccgag aatagccccg ctcggctcct 2220
gcatgcgtaa gagccagtcc attccccatt tgacttcgtc aagcaggtcg gggacaccgt 2280
tgccggattc cgggatgcca aaatcatcgg taaagacgtc aggccgccct tgataggcaa 2340
ggagcagctc caggatgacg cgccccgtcc actcgctgta cttgttgaaa tcgcccgcat 2400
cgaaccaacc gccgctgaga tcgcgctcca aggaggcatt ccccatatcc cagatggggc 2460
ggctggcgac gtcctgcggg tgagaagcgg catcggccca gttcgcgtgg gcgtagggca 2520
cctccttggc aaacccggag cgctgataga agaacatgcg cacggcctcg cgcaggacaa 2580
catcgtaaac atccgcgcca atggcgaaac tatcggaatg agtgttgttg gcaggatcgt 2640
ggatgcggta gtggccgggc tcggcaacta ccgtaaaatc aaaccaccac acgcggtctc 2700
ccgattgaat atggatggcg ccgccgttcc acgggaccgg tgagccggag aaaaccacga 2760
cgccatcgtt cacgcgacgg acctccagcg ttgcgccggg gctgtagctc tcggcgctgt 2820
tccagccaat ctgcgggtcg gcgatcaccg ccaccttggt ggcatcggcg gggtaaccga 2880
attggtcgat gcggatttta tcggtgtggg tggaggcgac gagggcggag ctgcccatga 2940
gcagcaagaa aaagcccgct gtcggcccga taccaaaaaa acgaataggg agagaaaaat 3000
tcatagcagg atgtggatac ggaaaggggg aaaacggtgc aaagacccaa gcccaacgct 3060
tggcgaaaac tggatggttg gtttatcaag aaaagcgctt ttgagccaaa agctgcgggc 3120
aatccttatt gcgtttcaca atattttcac atcgtcggcg gcacgacttt tcgatgggcg 3180
acttgacagc gtattctctc aggcgcgagg ctgcaaacct tatgaaaaaa ggcccgcgca 3240
gcgatctgtc cccggtcaaa atccagtcaa ggtttgttca agggtttgag gtctgataga 3300
ggcacagtcg agccatcagc agtcgcattg agtagggttg ttggagaaag tgtgcaaatg 3360
accgctgccg aaggaactgt ggagacaaaa agcatatttt cctcgccaag 3410

<210> SEQ ID NO: 2
<211> LENGTH: 1746
<212> TYPE: DNA
<213> ORGANISM: Bacillus sp.
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(1746)
<223> OTHER INFORMATION: isolated from environmental sample from Sonachi Lake, Kenya
<400> SEQUENCE: 2
atgaattttt ctctccctat tcgtttttttt ggtatcgggc cgacagcggg ctttttcttg 60
ctgctcatgg gcagctccgc cctcgtcgcc tccacccaca ccgataaaat ccgcatcgac 120
caattcggtt accccgccga tgccaccaag gtggcggtga tcgccgaccc gcagattggc 180
tggaacagcg ccgagagcta cagccccggc gcaacgctgg aggtccgtcg cgtgaacgat 240

RAW SEQUENCE LISTING DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:22

Input Set : **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

ggcgtcgtgg ttttctccgg ctcaccggtc ccgtggaacg gcggcgccat ccatattcaa 300
tcgggagacc gcgtgtggtg gtttgatttt acggtagttg ccgagcccgg ccactaccgc 360
atccacgatc ctgccaacaa cactcattcc gatagtttcg ccattggcgc ggatgtttac 420
gatgttgtcc tgcgcgaggc cgtgcgcatg ttcttctatc agcgctccgg gtttgccaag 480
gaggtgccct acgcccacgc gaactgggcc gatgccgctt ctcacccgca ggacgtcgcc 540
agccgcccca tctgggatat ggggaatgcc tccttggagc gcgatctcag cggcggttgg 600
ttcgatgcgg gcgatttcaa caagtacagc gagtggacgg ggcgcgtcat cctggagctg 660
ctccttgcct atcaagggcg gcctgacgtc tttaccgatg attttggcat cccggaatcc 720
ggcaacggtg tccccgacct gcttgacgaa gtcaaatggg gaatggactg gctcttacgc 780
atgcaggagc cgagcggggc tattctcggg aaagtttccg tgacggggca ccagagcgcc 840
agcccgccga gcaccgacac ccatccgcgt tactacggcc ccgtctcgac cgaggccact 900
gccatggctg ccgccgcttt cgccctcggg gcgactgtct ttgagagcgt aggcatgagc 960
gattatgccg tcaccctcga atccgctgcc atcgccgcat ggaattggac catggtccat 1020
cctcatgtgc ctttcgacaa caccggattt gcttccgtga gcccctcgcg caatgcccac 1080
gatacgctag ccaaccgtgt gatggccgcc gccatgctct tcgagcgcac gggcggtgcg 1140
gtctatcgcg atttcttcga tgtgcgctat ctcgatatgg agcccgtgca gtggtggtat 1200
ttcttccgct ttcaaggtga gctacaaaaa gctctcgccc actacacgac cctcccggga 1260
gcgacgccta gcgtctctgc cgacatccgc aaccggatgg cggcttctat aaacggtggg 1320
gaqtttctcg gtgcatggaa taaccagacc gacgcctacc gcgcctacct gaaggatcag 1380
gattacacgt ggggctcgaa caaaacaaag tctcaggctg gctttttctt cgagggggta 1440
cggaggcttg ggctcaaccc cgccgacgcg gccgcgcatc gcgatgccgc gatgggctac 1500
ctgcactatc tccacggagt gaacccaatg ggcatggtct atcttagcaa catgtatgcc 1560
agcggcgctg accgtgcagc caatgaaatc taccaccatt ggttccgcga tggccggact 1620
gggacaatgc cctcacttca ctctacggtc ccgctcctgg ttttctttcg ggcgggtccc 1680
aatgcccaaa tacagcggaa gcattcaggc gatccgagac caacccgtgc aaaaagccta 1740
ccttga 1746

<210> SEQ ID NO: 3
<211> LENGTH: 581
<212> TYPE: PRT
<213> ORGANISM: Bacillus sp.
<220> FEATURE:
<221> NAME/KEY: VARIANT
<222> LOCATION: (1)...(581)
<223> OTHER INFORMATION: isolated from environmental sample from Sonachi Lake, Kenya
<400> SEQUENCE: 3

```
Met Asn Phe Ser Leu Pro Ile Arg Phe Phe Gly Ile Gly Pro Thr Ala
 1               5                  10                  15
Gly Phe Phe Leu Leu Leu Met Gly Ser Ser Ala Leu Val Ala Ser Thr
            20                  25                  30
His Thr Asp Lys Ile Arg Ile Asp Gln Phe Gly Tyr Pro Ala Asp Ala
        35                  40                  45
Thr Lys Val Ala Val Ile Ala Asp Pro Gln Ile Gly Trp Asn Ser Ala
    50                  55                  60
Glu Ser Tyr Ser Pro Gly Ala Thr Leu Glu Val Arg Arg Val Asn Asp
65                  70                  75                  80
Gly Val Val Val Phe Ser Gly Ser Pro Val Pro Trp Asn Gly Gly Ala
                85                  90                  95
Ile His Ile Gln Ser Gly Asp Arg Val Trp Trp Phe Asp Phe Thr Val
            100                 105                 110
```

RAW SEQUENCE LISTING DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:22

Input Set : **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

```
161 Val Ala Glu Pro Gly His Tyr Arg Ile His Asp Pro Ala Asn Asn Thr
162         115                 120                 125
163 His Ser Asp Ser Phe Ala Ile Gly Ala Asp Val Tyr Asp Val Val Leu
164     130                 135                 140
165 Arg Glu Ala Val Arg Met Phe Phe Tyr Gln Arg Ser Gly Phe Ala Lys
166 145                 150                 155                 160
167 Glu Val Pro Tyr Ala His Ala Asn Trp Ala Asp Ala Ala Ser His Pro
168                 165                 170                 175
169 Gln Asp Val Ala Ser Arg Pro Ile Trp Asp Met Gly Asn Ala Ser Leu
170             180                 185                 190
171 Glu Arg Asp Leu Ser Gly Gly Trp Phe Asp Ala Gly Asp Phe Asn Lys
172         195                 200                 205
173 Tyr Ser Glu Trp Thr Gly Arg Val Ile Leu Glu Leu Leu Leu Ala Tyr
174     210                 215                 220
175 Gln Gly Arg Pro Asp Val Phe Thr Asp Asp Phe Gly Ile Pro Glu Ser
176 225                 230                 235                 240
177 Gly Asn Gly Val Pro Asp Leu Leu Asp Glu Val Lys Trp Gly Met Asp
178                 245                 250                 255
179 Trp Leu Leu Arg Met Gln Glu Pro Ser Gly Ala Ile Leu Gly Lys Val
180             260                 265                 270
181 Ser Val Thr Gly His Gln Ser Ala Ser Pro Pro Ser Thr Asp Thr His
182         275                 280                 285
183 Pro Arg Tyr Tyr Gly Pro Val Ser Thr Glu Ala Thr Ala Met Ala Ala
184     290                 295                 300
185 Ala Ala Phe Ala Leu Gly Ala Thr Val Phe Glu Ser Val Gly Met Ser
186 305                 310                 315                 320
187 Asp Tyr Ala Val Thr Leu Glu Ser Ala Ala Ile Ala Ala Trp Asn Trp
188                 325                 330                 335
189 Thr Met Val His Pro His Val Pro Phe Asp Asn Thr Gly Phe Ala Ser
190             340                 345                 350
191 Val Ser Pro Ser Arg Asn Ala His Asp Thr Leu Ala Asn Arg Val Met
192         355                 360                 365
193 Ala Ala Ala Met Leu Phe Glu Arg Thr Gly Gly Ala Val Tyr Arg Asp
194     370                 375                 380
195 Phe Phe Asp Val Arg Tyr Leu Asp Met Glu Pro Val Gln Trp Trp Tyr
196 385                 390                 395                 400
197 Phe Phe Pro Phe Gln Gly Glu Leu Gln Lys Ala Leu Ala His Tyr Thr
198                 405                 410                 415
199 Thr Leu Pro Gly Ala Thr Pro Ser Val Ser Ala Asp Ile Arg Asn Arg
200             420                 425                 430
201 Met Ala Ala Ser Ile Asn Gly Gly Glu Phe Leu Gly Ala Trp Asn Asn
202         435                 440                 445
203 Gln Thr Asp Ala Tyr Arg Ala Tyr Leu Lys Asp Gln Asp Tyr Thr Trp
204     450                 455                 460
205 Gly Ser Asn Lys Thr Lys Ser Gln Ala Gly Phe Phe Phe Glu Gly Val
206 465                 470                 475                 480
207 Arg Arg Leu Gly Leu Asn Pro Ala Asp Ala Ala Ala His Arg Asp Ala
208                 485                 490                 495
209 Ala Met Gly Tyr Leu His Tyr Leu His Gly Val Asn Pro Met Gly Met
```

RAW SEQUENCE LISTING DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:22

Input Set : **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

```
210                 500                 505                 510
211 Val Tyr Leu Ser Asn Met Tyr Ala Ser Gly Ala Asp Arg Ala Ala Asn
212         515                 520                 525
213 Glu Ile Tyr His His Trp Phe Arg Asp Gly Arg Thr Gly Thr Met Pro
214     530                 535                 540
215 Ser Leu His Ser Thr Val Pro Leu Leu Val Phe Phe Arg Ala Gly Pro
216 545                 550                 555                 560
217 Asn Ala Gln Ile Gln Arg Lys His Ser Gly Asp Pro Arg Pro Thr Arg
218                 565                 570                 575
219 Ala Lys Ser Leu Pro
220             580
```

VERIFICATION SUMMARY DATE: 04/05/2007
PATENT APPLICATION: **US/10/549,943** TIME: 07:50:23

Input Set : **A:\796_2_PCT_SeqListing.TXT**
Output Set: **N:\CRF4\04052007\J549943.raw**

L:15 M:270 C: Current Application Number differs, Replaced Current Application Number
L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date